BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68

A1

(11) International Publication Number: WO 97/21837

(43) International Publication Date: 19 June 1997 (19.06.97)

(21) International Application Number: PCT/US96/20075

(22) International Filing Date: 13 December 1996 (13.12.96)

(30) Priority Data:
60/008,673 15 December 1995 (15.12.95) US

(71) Applicant: AMERSHAM LIFE SCIENCE, INC. [US/US]; 26111 Miles Road, Cleveland, OH 44128 (US).

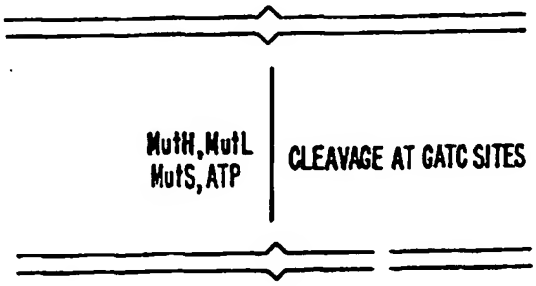
(72) Inventors: MODRICH, Paul, L.; 5820 Brisbane Drive, Chapel Hill, NC 27514 (US). SMITH, Jane, E.; 404 West Carver Street, Durham, NC 27704 (US).

(74) Agents: WARBURG, Richard, J. et al.; Lyon & Lyon, First Interstate World Center, Suite 4700, 633 West Fifth Street, Los Angeles, CA 90071-2006 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHODS FOR THE DETECTION AND REMOVAL OF MUTANT SEQUENCES THAT ARISE DURING ENZYMATIC AMPLIFICATION USING MISMATCH REPAIR SYSTEMS



(57) Abstract

Method for removing DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprising the steps of denaturing and reannealing the population of DNA duplexes, contacting the reannealled DNA duplexes with a mismatch repair system such that each strand is cleaved in DNA duplexes containing a base pair mismatch, and separating the cleaved DNA duplexes from uncleaved DNA duplexes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

DESCRIPTION

METHODS FOR THE DETECTION AND REMOVAL OF MUTANT SEQUENCES THAT ARISE DURING ENZYMATIC AMPLIFICATION USING MISMATCH REPAIR SYSTEMS.

FIELD OF THE INVENTION

The present invention relates to methods for the detection and removal of mutant sequences in DNA molecules that are the product of enzymatic amplification.

BACKGROUND OF THE INVENTION

The following is a discussion of the relevant art, none of which is admitted to be prior art to the appended claims.

DNA polymerase errors occurring during PCR amplification result in the presence of mutations in the amplified product. This problem can be particularly acute with *Taq* DNA polymerase which lacks a proofreading exonuclease and has a base substitution error rate on the order of $1/10^4$ to $1/10^5$ nucleotides polymerized under PCR conditions (Eckert and Kunkel, Nucleic Acids Res., 18:3739, 1990; Mattila et al.,Nucleic Acids Res., 19:4967, 1991; Saiki et al., Science, 239:487, 1988; Tindall and Kunkel, Biochemistry, 27:6008, 1988). The significance of error rates of this magnitude has been pointed out by Keohavong and Thilly (Keohavong and Thilly, Proc. Natl. Acad. Sci.

U.S.A., 86:9253, 1989), who noted that at a misincorporation rate of $2/10^4$, 10^6 -fold (twenty cycle) amplification of a 100 base pair sequence yields a population of product molecules, each of which has an 80% probability of containing a mutation somewhere in its sequence (Keohavong and Thilly, supra). The frequency of polymerase errors during PCR can be estimated from equations 1 and 6 of Luria and Delbrück (Luria, S.E. & Delbrück, M. Genetics 28:491-511, 1943) as

$$f = 2\ell Na \quad \text{(Equation 1)}$$

where f is the expected fraction of product molecules that contain a mutation somewhere in their sequence, ℓ is the length of the amplified segment in bp, N is the number of cycles, and a is the error rate for the polymerase expressed per nucleotide incorporated. This problem has been alleviated to some extent by identification of the thermostable *Pfu* and *Tli* (Vent™) DNA polymerases, which have proofreading activity and display a two to ten-fold improvement in fidelity relative to *Taq* (Lundberg et al., Gene, 108:1, 1991; Mattila et al., supra). However, given that the probability of polymerase misincorporation event per cycle is also proportional to the size of the sequence being amplified, polymerase-generated mutations remain a significant problem for extensive amplification of sequences in the kilobase range.

SUMMARY OF THE INVENTION

The current invention concerns detection and removal of mutant sequences in DNA molecules that are the product of enzymatic amplification. It is based on the use of mismatch repair reactions, such as the MutHLS reaction of E. coli which is responsible for the initiation of bacterial methyl-directed mismatch repair. The presence of a mismatch within DNA molecules provokes cleavage at a GATC sequence located in the vicinity of the mispair in a reaction that depends on MutH, MutL, MutS and ATP (Au et al., J. Biol. Chem., 267:12142, 1992). Hemimethylated GATC is incised on the unmethylated strand. Heteroduplex DNAs devoid of GATC methylation have been reported to be subject to mismatch-provoked single and double strand cleavage at such sites with the latter reaction evident after prolonged incubation or elevated concentrations of MutH, MutL, and MutS proteins (Au et al., supra) (See Fig. 1). In the current invention single or double strand cleavage is utilized to detect the presence of mutants that are the result of enzymatic amplification. The double strand cleavage reaction can also be exploited to remove mutant sequences from a population of amplified molecules.

The present invention is applicable to enzymatically amplified populations of DNA molecules. Such methods include polymerase chain reaction (PCR), and reverse transcription/ polymerase chain reaction (RT/PCR).

If the current invention is to be utilized for the removal of mutations that are the result of enzymatic

enzymes, such as MutHLS, cleaved and uncleaved DNA duplexes are separated, e.g., by gel electrophoresis, and the relative amount of DNA duplexes present in each fraction is determined by standard techniques used for DNA quantitation. The extent of MutHLS cleavage at this point provides a direct estimate of the fraction of PCR product molecules that contain one or more polymerase-induced mutations.

The methods of the claimed invention require the presence of a sequence that is subject to mismatch provoked endonucleolytic cleavage, such as a dGATC site, in the sequence to be analyzed, so that endonucleolytic cleavage can occur. If such a site does not exists in the region that is amplified a sequence that is subject to mismatch provoked endonucleolytic cleavage can be introduced by using a primer that contains this sequence. Thus, the invention also features a method to detect or remove errors in sequences that do not contain a sequence that is subject to mismatch provoked endonucleolytic cleavage. Molecules without dGATC sites may be screened or removed with MutHLS cleavage by introducing a dGATC site into the primer used for amplification. Applicant has found that dGATC sites 50-100 bp from the end of a molecules are sufficient for mutation screening. Sequences other than GATC, that are subject to mismatch provoked endonucleolytic cleavage by mismatch repair enzymes from other organisms, can also be utilized in the present invention.

By "denaturing and reannealing" is meant methods known to those who practice the art, by which the hydrogen bonding of DNA duplex molecules is sequentially disrupted and then allowed to reform. Preferred methods for denaturing and reannealing include elevation of temperature followed by lowering of temperature and elevation of pH (e.g., pH 13) followed by neutralization and annealing at an appropriate temperature. Alternative methods include use of enzymes such as RNase H in RT/PCR methodology.

By "mismatch repair system" is meant proteins that include a GATC endonuclease such E. coli MutH or a protein functionally homologous to E. coli MutH (the protein may cleave at a site different from GATC), a mispair recognition protein such E. coli MutS or a protein functionally homologous to E. coli MutS, and proteins that participate in the activation of the GATC endonuclease such as E. coli MutL or a protein functionally homologous to E. coli MutL, and necessary cofactors such as $MgCl_2$ and ATP.

By "separating" is meant isolating the cleaved molecules from uncleaved molecules by methods that achieve physical separation, such as electrophoresis or HPLC.

In preferred embodiments, the method of separation is by electrophoresis through a gel; the mismatch repair system comprises components of the methyl-directed mismatch repair system of E. coli and includes the Muts, MutL, and MutH proteins.

By "electrophoresis through a gel" is meant the method of size fractionation by electrophoretic mobility, a

Other organisms, including man, are known to possess systems for recognition and repair of DNA mispairs, which, as one skilled in the art would appreciate, comprise proteins functionally homologous to the MutHLS proteins of E. coli, which may be used in the present invention.

Thus in a first aspect the invention features a method for removing DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes. The method comprises denaturing and reannealing the population of DNA duplexes, contacting the reannealed DNA duplexes with a mismatch repair system such that each strand is cleaved in DNA duplexes containing a base pair mismatch, and separating said cleaved DNA duplexes from uncleaved DNA duplexes.

By "polymerase generated mutations" is meant a misincorporation of a nucleotide by a DNA polymerase during the course of DNA amplification, so as to produce incorrect pairing between the bases of two nucleotides located on complementary strands of DNA, i.e., base pairs that are not A:T or G:C, or the presence of 1, 2 or 3 extra unpaired, adjacent nucleotides on one strand (an insertion/deletion mismatch).

By "enzymatically amplified DNA duplexes" is meant DNA that has been amplified by an enzymatic amplification reaction. Examples of such reactions include the polymerase chain reaction and reactions utilizing reverse transcription and subsequent DNA amplification of one or more expressed RNA sequences.

procedure well known by those skilled in the art. Gel electrophoresis can either be conventional or pulse-field.

The "components of the methyl-directed mismatch repair system of E. coli" includes the proteins MutH, MutL and MutS and cofactors $MgCl_2$ and ATP.

In a second aspect the invention features a method for removing DNA molecules containing one or more polymerase-generated mutations in a population of DNA duplexes that have been enzymatically amplified using primers containing 5'-hydroxyl termini. The method comprises denaturing and reannealing the population of DNA duplexes, contacting the reannealed DNA duplexes with a mismatch repair system such that each strand in DNA duplexes containing a base pair mismatch is cleaved so as to produce 5'-phosphate termini, and further contacting the population of reannealed DNA duplexes with exonucleases so that the DNA duplexes containing a base pair mismatch are enzymatically degraded.

The use of primers with 5'-hydroxyl termini prevent exonuclease attack at sites other than cleavage sites generated by the mismatch repair system. Primers are typically constructed with such 5'-hydroxyl termini.

By "exonucleases" is meant a combination of an exonuclease that preferentially degrades duplex DNA and also has a preference for 5'-phosphate, such as λ exonuclease, and a single strand specific exonuclease, such as exonuclease I.

By "enzymatically degraded" is meant broken down to single nucleotides or dinucleotides.

In a preferred embodiment, the mismatch repair system comprises components of the methyl-directed mismatch repair system of E. coli and includes the Muts, MutL, and MutH proteins.

In a third aspect the invention features a method for rending inert to further amplification DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes. The method comprises denaturing and reannealing the population of DNA duplexes, contacting the reannealed DNA duplexes with a mismatch repair system such that each strand is cleaved in DNA duplexes containing a base pair mismatch, and, further contacting the cleaved DNA duplexes with dideoxynucleoside-5'-triphosphate and a DNA polymerase capable of adding a dideoxynucleoside-5'-monophosphate moiety to 3'-termini produced by mismatch-provoked cleavage.

By "inert to further amplification" is meant unable to be subject to further exponential amplification as the cleaved DNA molecules have incorporated a chain terminating nucleotide incapable of supporting further elongation.

By "dideoxynucleoside-5'-triphosphate" is meant a nucleoside-5'-triphosphate with the 3'-hydroxyl group replaced by a hydrogen so as to result in a chain terminating analog. This analog will block the growth of a

new DNA chain as it lacks a 3'-hydroxyl necessary to form a phosphodiester bond.

In a preferred embodiment, the mismatch repair system comprises components of the methyl-directed mismatch repair system of E. coli and includes the Muts, MutL, and MutH proteins and the dideoxynucleoside-5'-triphosphate is dideoxyguanosine-5'-triphosphate.

In a fourth aspect the invention features a method for determining the fraction of an enzymatically amplified DNA population that contains polymerase-generated mutations. The method comprises denaturing and reannealing the population of DNA duplexes, contacting the reannealed DNA duplexes with a mismatch repair system such that at least one strand is cleaved in DNA duplexes containing a base pair mismatch, separating the cleaved DNA duplexes from uncleaved DNA duplexes, and determining the fraction of cleaved DNA duplexes relative to uncleaved DNA duplexes as an indication of the fraction of enzymatically amplified DNA that contain polymerase-generated mutations.

By "at least one strand is cleaved" is meant that an endonucleolytic incision is introduced into one or both strands of a DNA molecule containing a base pair mismatch. The length of incubation and the concentration of the mismatch repair enzymes determine whether one or two endonucleolytic incisions are introduced.

By "separating the cleaved DNA duplexes from uncleaved DNA duplexes" is meant the physical separation of

the two classes of molecules and includes electrophoresis under denaturing conditions.

By "determining the fraction of cleaved DNA duplexes relative to uncleaved DNA duplexes" is meant detecting and quantitating DNA, after separation of the cleaved and uncleaved molecules, so that the relative amounts in each fraction can be determined. Detection and quantitation of DNA can be carried out by methods familiar to those who practice the art, utilizing known labeling, staining, and quantitation techniques for nucleic acids.

In a preferred embodiment, the mismatch repair system comprises components of the methyl-directed mismatch repair system of E. coli and includes the Muts, MutL, and MutH proteins.

In a fifth aspect, the invention features a method for detecting the presence of DNA polymerase-generated mutations in a population of enzymatically amplified DNA duplexes. The method comprises denaturing and reannealing the population of DNA duplexes, contacting the reannealed DNA duplexes with a mismatch repair system under conditions such that a duplex containing a polymerase generated mutation is modified by the introduction of an endonucleolytic incision in at least one strand of the duplex, and detecting the product of the endonucleolytic incision as an indication of the presence of polymerase generated mutations.

The endonucleolytic incision or cleavage product can be detected by any method which detects a difference

between the incision product and an unmodified molecule. Such methods include those that detect differences in size or electrophoretic mobility.

In preferred embodiments, detection of the product of the endonucleolytic incision is by altered electrophoretic mobility under denaturing conditions; the mismatch repair system comprises components of the methyl-directed mismatch repair system of E. coli and includes the Muts, MutL, and MutH proteins.

By "altered electrophoretic mobility " is meant mobility on a gel that is different from or relative to an unmodified, i.e., uncleaved molecule.

Denaturing conditions are know to those who practice the art, and include the use of urea.

In a sixth aspect, the invention features a method for detecting the presence of DNA polymerase-generated mutations in a population of enzymatically amplified DNA duplexes produced from DNA duplexes lacking a sequence subject to mismatch provoked endonucleolytic cleavage. The method comprises the steps of enzymatically amplifying a population of DNA molecules utilizing primers containing a sequence subject to mismatch provoked endonucleolytic cleavage, denaturing and reannealing the population of DNA duplexes, contacting the reannealed DNA duplexes with a mismatch repair system under conditions such that an endonucleolytic incision is introduced in at least one strand of a duplex containing a polymerase generated mutation, and detecting the product of the endonucleolytic

incision as an indication of the presence of polymerase generated mutations.

By "primers that contain a sequence subject to mismatch provoked endonucleolytic cleavage" is meant primers engineered to contain a sequence that is specifically cleaved in response to the presence of a mismatch by components of a mismatch repair system. Primers with such sites can be constructed by standard cloning techniques known to those who practice the art.

The invention also features the use of primers containing a sequence subject to mismatch provoked endonucleolytic cleavage for the removal of DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes produced from DNA duplexes lacking a sequence subject to mismatch provoked endonucleolytic cleavage. The method comprises the steps of enzymatically amplifying a population of DNA molecules utilizing primers containing a sequence subject to mismatch provoked endonucleolytic cleavage, denaturing and reannealing the population of DNA duplexes, contacting the reannealed DNA duplexes with a mismatch repair system under conditions such that each strand is cleaved in a DNA duplex containing a polymerase generated mutation, and separating the cleaved DNA duplexes from uncleaved DNA duplexes.

In preferred embodiments, the sequence subject to endonucleolytic cleavage is a d(GATC) site; detection of the product of endonucleolytic incision is by altered

electrophoretic mobility under denaturing conditions; the mismatch repair system comprises components of the methyl-directed mismatch repair system of E. coli and includes the Muts, MutL, and MutH proteins.

In a seventh aspect, the invention features various kits for detecting, removing, or rendering inert to further amplification DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes. One such kit useful for amplifying DNA molecules and for removing DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprises primers with 5'-hydroxyl termini, components of a mismatch repair system, and exonucleases. Another kit for removing DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprises components of a mismatch repair system and exonucleases. The invention also features a kit for rendering inert to further amplification DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprising components of a mismatch repair system, a dideoxynucleoside-5'-triphosphate, and a DNA polymerase. Another kit is useful for amplifying DNA molecules lacking a sequence subject to mismatch provoked endonucleolytic cleavage and for detecting or removing DNA molecules containing one or more polymerase-generated mutations from this population of enzymatically amplified DNA duplexes.

This kit comprises primers containing sequence subject to mismatch provoked endonucleolytic cleavage, and components of a mismatch repair system.

The present invention offers several advantages over other methods used to detect and/or eliminate polymerase-induced mutations in amplification reactions.

The reliability and applicability of molecular methods for mutation detection have been subject to technical limitations, high levels of background signal for perfectly paired DNA sequences and failure to detect all mutations. Methods relying solely upon differential resolution of DNA fragments in polyacrylamide gels are subject to severe size constraints. Chemical approaches to mutation detection are subject to background reactivity with perfectly paired sequences. Enzymatic methods have proven less robust in their sensitivity to different mutations and in some cases are subject to background signals with perfectly paired DNA. Mismatch repair systems, such as the MutHLS-dependent d(GATC) cleavage reaction, circumvent many of these limitations, as they are exquisitively sensitive to the presence of mismatches and have a high degree of accuracy.

Polymerases with high efficiency and/or proof reading capacity are limited in their ability to reduce the number of polymerase-generated errors as the probability of a polymerase misincorporation event increases as does the number of amplification cycles and is proportional to the size of the sequence being amplified. In distinction, the

methods of the present invention are not subject to these limitations as they are utilized after the amplification event. In addition, as these methods are based on the use of extremely sensitive and accurate mispair recognition proteins, most errors should be detected and eliminated, regardless of the number of cycles or the size of the amplified sequence. The notable exception is C-C mismatches, which are not recognized by the E. coli methyl-directed system.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

BRIEF DESCRIPTION OF THE FIGURES

Fig.1 is a schematic representation of the MutHLS double strand cleavage reaction.

Fig.2 is a schematic representation of denaturation and reannealing of an enzymatically applied population followed by the MutHLS double strand cleavage reaction of molecules containing a base pair mismatch.

Fig. 3A is a graphic representation of the results of MutHLS cleavage of 1169 bp phage f1 gene VII homohybrids obtained after 10, 20 or 30 cycles of PCR amplification using Pfu polymerase. The x-axis represents number of PCR cycles. The y-axis represents homohybrids cleaved (%). Error bars indicate the standard error for four independent experiments.

Fig. 3B represents the results of MutHLS cleavage of 1360 bp lacI homohybrids obtained after 25 cycles of amplification using Taq, Vent, or Pfu polymerases. The x-axis indicates the type of polymerase utilized. The y-axis represents homohybrids cleaved (%). Error bars indicate the standard error for four independent experiments.

Fig. 4 is a graphic representation of the dependence of MutHLS cleavage on dNTP pool composition and polymerase used during amplification. The x-axis indicates the dGTP pool bias. The y-axis represents homohybrids cleaved (%). Polymerases utilized were: Pfu (◆), Vent (●) and Taq (■).

Fig. 5A is a schematic representation of a heterohybrid containing a base pair mismatch and an introduced dGATC site. Location of the single nucleotide insertion/deletion mutation is indicated as is the d(GATC) site, which are separated by about 1,000 bp.

Fig. 5B is a graphic representation of the dependence of the efficiency of MutHLS cleavage of a heterohybrid on the distance between the dGATC site and the proximal DNA end. The x-axis represents distance in base pairs. The y-axis represents d(GATC)cleavage (%).

Fig. 6 is a schematic representation of the experimental design for assessing the efficiency of MutHLS treatment for removal of sequences containing polymerase-induced mutations that occur during the process of PCR.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The current invention encompasses methods for the detection and removal of mutations that are a result of polymerase errors that occur during enzymatic amplification of nucleic acids. The methods are based on the use of components of mismatch repair systems. The components and use of such systems is extensively described in "Methods of Analysis and Manipulation of DNA Utilizing Mismatch Repair Systems", WO 95/12688, which is incorporated herein, in its entirety, by reference. In most instances, sequences will be amplified by techniques familiar to those who practice the art, prior to application of the claimed methods.

PCR amplification of DNA

PCR reactions were carried out to generate substrates for use in experiments described in Examples 1-3. The PCR primers utilized are indicated in the specific examples. Unless otherwise noted, reactions (100 ul) contained 20 mM Tris-HCl (pH 8.2), 10mM KCl, 6mM $(NH_4)_2SO_4$, 4mM $MgCl_2$, 0.1% Triton X-100, 10μg/ml bovine serum albumin (BSA), 1mM of each deoxyribonucleoside-5'-triphosphate (dNTP) (Pharmacia Biotech), 100 pmol of each primer, 5μg T4 gene 32 protein (Boehringer Mannheim), 100 ng template DNA and 2.5 units of native Pfu DNA polymerase (Stratagene). Reactions in which synthetic products were uniformly labeled also contained 70 μCi of [α^{32}P]dTTP (3000 Ci/mmol, DuPont/New England Nuclear). Reactions in which synthetic products were end-labeled contained 100 pmol of the

appropriate primer labeled with T4 polynucleotide kinase (Amersham) and [γ^{32}]ATP (3000 Ci/mmol, DuPont/New England Nuclear) as described (Sambrook et al., Molecular Cloning a Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY , 1989). PCR reactions (15 cycles) were performed using a Perkin Elmer Gene Amp 96000 thermocycler with incubations at 94°C for 15 sec, 60°C for 15 sec and 72°C for 90 sec, 3 min, 4 min or 6 min for amplification of 400 bp, 1.3 kb, 1.7kb and 2.5 kb sequences respectively.

PCR reactions, in which Pfu, Vent (New England Biolabs) and Taq (Amersham) polymerases were compared, using buffer conditions recommended by the manufacturer. Reactions contained 1X buffer supplied with each polymerase as well as 200 μM of each dNTP, 100 pmol of each primer, 5μg T4 gene 32 protein, 15 ng template DNA and 2.5 units of polymerase. The volume of each reaction was 100 μl. Reactions in which DNA was uniformly labeled contained 80 μCi of [α^{32}P]dTTP. Reactions proceeded for 25 cycles with each cycle consisting of 15 sec at 94°C, 15 sec at 55°C and 30 sec at 72°C.

To avoid introduction of contaminating DNA into PCR reactions, buffer components were made fresh daily and reactions were assembled in a laminar flow hood using filtered pipette tips. Products were extracted with phenol and ether, precipitated with ethanol and quantitated by an ethidium bromide dot method. Samples (0.5ml of an appropriate dilution) and DNAs of known concentration were

added to 8 μl of 1 μg/ml ethidium bromide and spotted onto plastic wrap. Ultraviolet-induced fluorescence was measured using a Photometrics cooled CCD imager. The concentration of PCR products was determined by comparison to the fluorescence of the standards.

MutHLS reactions

MutHLS reactions were carried out as follows in experiments described in Examples 1-3. Denaturation/reannealing reactions (20μl) contained 2.5 μg of unlabeled PCR product, 0.5 μg of uniformly ^{32}P labeled PCR product, 10 mM NaCL, 1 mM EDTA and 50 mM Hepes-KOH (pH 8.0). Freshly prepared 10 N Na OH (0.6μl) was added to a final concentration of 300 mM and the mixture was incubated at room temperature for five min. The solution was neutralized by addition of acetic acid to a final concentration of 300 mM, KCl to 100 mM and potassium phosphate (pH 7.4) to 100 mM, and the DNA hybridized at 65°C for 30 min followed by 30 min at 37°C. Reactions were then bound to a silica matrix spin column (Pierce Xtreme DNA purification columns) and eluted with dH_2O to remove PCR primers, dNTPs and salts.

Reactions (10 μl) (Au et al., supra) contained 50 mM Hepes-KOH (pH 8.0), 20 mM KCl, 4 mM $MgCl_2$, 1mM dithiothreitol (DTT), 50 μg/ml BSA, 2 mM ATP, approximately 10,000 cpm of PCR DNA (50-200 ng), 250 ng MutS (Su et al., Proc. Natl. Acad. Sci. USA, 83:5057, 1986), 600 ng MutL (Grilley et al., J. Biol. Chem., 264:1000, 1989) and 0.9 ng

MutH (Welsh et al., J. Biol. Chem., 262:15624, 1987). DNA and buffer components were preincubated at 37°C for 8 min, reactions initiated by adding a premixed solution of MutH, MutL and MutS and incubation continued for 15 min at 37°C to produce single strand cleavage products. After addition of 0.5 μl of 0.5 M EDTA and 20 μl of deionized formamide containing 0.05% bromophenol blue and 0.05% xylene cyanol, DNA products were analyzed by electrophoresis through 6% polyacrylamide in 89 mM Tris, 89 mM boric acid, 2 mM EDTA (final pH of 8.5) and 8 M urea. DNA species were visualized by autoradiography and quantitated using a Molecular Dynamics Phosphorimager.

Example 1: Polymerase errors during amplification account for d(GATC) cleavage of homohybrid products

To determine whether the observed MutS-dependent d(GATC) cleavage of homohybrids was due to damage incurred during DNA preparation or to genetic variation introduced during PCR amplification, the dependence of the level of such cleavage on PCR reaction conditions was examined.

Phage f1 gene VII sequences of 1169 bp were amplified from the plasmid templates of Ivey-Hoyle et al. (Ivey-Hoyle et al., J. Mol. Biol., 224:1039, 1992)for 10, 20 or 30 cycles using Pfu polymerase. Wild type *lacI* sequences 1360 bp in length were amplified from the plasmid clones of Matteson et al.(Matteson et al. Nucleic Acids Res., 19:3499, 1991) for 25 cycles using Pfu, Vent or Taq

polymerases. Both *lacI* and phage f1 gene VII sequences were amplified utilizing as the forward primer CAGAACTTTAAAAGTGCTCAT (SEQ. ID. NO. 1) and as the reverse primer ATGCAGCAACGAGACGTCACG (SEQ. ID. NO. 2). PCR products consisted of the gene fragments of interest as well as surrounding vector sequence. Both phage f1 gene VII sequences and *lacI* homohybrid molecules were prepared by a denaturation and reannealing step.

The fraction of amplified phage f1 gene VII homohybrids cleaved by MutH increased with the number of cycles (See Fig. 3A), a finding consistent with either cycle-dependent DNA damage or polymerase-induced mutations. However, the degree of homohybrid cleavage was also found to depend on the polymerase used for PCR amplification. Thus, cleavage of amplified lacI homohybrids was highest when *Taq* polymerase was used for amplification, intermediate with Vent polymerase and lowest with Pfu polymerase (See Fig. 3B). These results parallel the error rates for these enzymes, with the lower fidelity of *Taq* polymerase due to absence of a 3' to 5' editing exonuclease (Lundberg et al., supra; Mattila et al, supra; Tindall et al., supra). Although a low level of template damage associated with thermal cycling cannot be excluded, these findings indicate that the majority of the homohybrid background signal is due to polymerase errors occurring during amplification.

Example 2: Use of dNTP pool bias during PCR amplification to determine detectability of nucleotide-substitution errors

A dNTP pool imbalance leads to an increased error rate during *in vitro* synthesis by DNA polymerases (Kunkel et al., J. Biol. Chem., 254:5718, 1979; Fersht, Proc. Natl. Acad. Sci. USA, 76:4946, 1979). This observation was exploited to test the utility of the MutHLS reaction for detection of PCR errors. For these experiments wild type lacI sequences 1360 bp in length were amplified for 15 cycles using Pfu, Vent or Taq polymerases under conditions of dGTP pool imbalance. Wild type *lacI* sequences were amplified from the plasmid colones of Matteson et al., (Matteson et al., Nucleic Acids Res., 19:3499, 1991) utilizing as the forward primer CAGAACTTTAAAAGTGCTCAT (SEQ. ID. NO. 1) and as the reverse primer ATGCAGCAACGAGACGTCACG (SEQ. ID. NO. 2). PCR products consisted of the gene fragments of interest as well as surrounding vector sequence. Under equimolar conditions, each dNTP was present at 1 mM. The concentration of dGTP was 2 mM in all other reactions, and concentrations of the other three dNTPs were 667μM 200 μM and 20 μM each. PCR products were denatured and reannealed, subjected to MutHLS cleavage, and products analyzed as previously described. As shown in Fig. 4, d(GATC) cleavage of homohybrids was dependent on the dGTP concentration bias. Homohybrids derived from amplification using Taq polymerase were subject to MutHLS-dependent d(GATC) cleavage to a greater degree than

homohybrids amplified under the same conditions using Pfu and Vent polymerases. Negligible PCR product was obtained in reactions using Taq polymerase in which dGTP was present in a 100-fold molar excess over the other dNTPs. Since the enzyme lacks a 3' exonuclease activity, high levels of misincorporated dNTPs induce chain termination (Innis et al., Proc. Natl. Acad. Sci. USA, 85:9436, 1988). Likewise homohybrids amplified using Vent polymerase were cleaved to a greater extent than homohybrids amplified using Pfu polymerase (See Fig. 4). Polymerase misincorporation errors are therefore readily detectable by MutS-dependent d(GATC) cleavage.

Example 3: Dependence of the efficiency of cleavage by activated MutH on the distance between a d(GATC) site and the end of a DNA heterohybrid

Although highly sensitive to mismatched base pairs, the MutHLS reaction can only be used for mutation screen if the sequence of interest contains a d(GATC) site. To determine the feasibility of introducing a d(GATC) site into a PCR primer in order to screen sequences lacking such a site, we have evaluated the dependence of the reaction on the distance of a d(GATC) site from a DNA end (See Fig. 5A). Heterohybrids were prepared after amplification of the replicative form of phage f1MR21 and f1MR22 (which contains one extra nucleotide relative to phage f1MR21) for 15 cycles using nested reverse PCR primers

GATAAGAGGTCATTTTTGCGG (SEQ. ID. NO. 3) (1470 bp PCR product); AGACCGGAAGCAAACTCCAAC (SEQ. ID. NO. 4) (1528 bP PCR product); GCCCGAAAGACTTCAAATATC (SEQ. ID. NO. 5) (1578 bp PCR product); TTATAGTCAGAAGCAAAGCGG (SEQ. ID. NO. 6) (1624 bp PCR product); GGATAGCGTCCAATACTGCGG (SEQ. ID. NO. 7) (1743 bp PCR product); ATCATAACCCTCGTTTACCAG (SEQ. ID. NO. 8) (1845 bp PCR product) and the same forward primer CCAGCAAGGCCGATAGTTTGA (SEQ. ID. NO. 9). Phage f1MR22 was constructed by the insertion of a synthetic oligonucleotide duplex (Parsons et al., Cell, 75:1227, 1993) into the replicative form of phage f1MR1 (Su et al., supra). PCR products were amplified, and heterohybrids were prepared by denaturing and reannealing a mixture of PCR products obtained from the mutant and wild type DNA. Heterohybrids were subject to MutHLS reaction and analyzed as previously described. Maximum cleavage observed with these heterohybrids was 20%, perhaps reflecting the large (1,000 bp) distance between the mutation and the d(GATC) site. As shown in Fig. 5B, the efficiency of mismatch-provoked cleavage increased with increasing distance of the d(GATC) site from the proximal end in the range of 50-150 bp reaching a maximum at the latter distance. These results suggest that a PCR primer with a d(GATC) site 50-100 nucleotides from an end would prove sufficient for the purpose of amplification and subsequent mutation screen using MutH, MutL and MutS.

Example 4: Removal of molecules containing a polymerase-generated mutation utilizing MutHLS reaction followed by gel electrophoresis

A population of molecules that have been enzymatically amplified is subject to denaturation/reannealing reactions. A solution of amplified product (20μl) is made to have a concentration of 10 mM NaCL, 1 mM EDTA and 50 mM Hepes-KOH (pH 8.0). Freshly prepared 10 N NaOH is added to a final concentration of 300 mM and the mixture is incubated at room temperature for five minutes. The solution is neutralized by addition of acetic acid to a final concentration of 300 mM, KCl to 100 mM and potassium phosphate (pH 7.4) to 100 mM, and the DNA hybridized at 65° C for 30 minutes followed by 30 minutes at 37°C. Reactions are then bound to a silica matrix spin column (Pierce Xtreme DNA purification columns) and eluted with dH20 to remove PCR primers, dNTPs and salts.

MutHLS reactions are carried out (10 μl) or adjusted for volume as necessary containing 50 mM Hepes-KOH (pH 8.0), 20 mM KCl, 4 mM $MgCl_2$, 1mM dithiothreitol (DTT), 50 mg/ml BSA, 2 mM ATP, PCR DNA (50-200 ng), 500 ng MutS, 1200 ng MutL and 1.8 ng MutH. DNA and buffer components are preincubated at 37°C for 8 minutes, reactions are initiated by adding a premixed solution of MutH, MutL, and MutS and incubation continued for 45 minutes at 37°C. Reactions are supplemented with additional MutS (500 ng), MutL (1200 ng) and MutH (1.8 ng) and incubated at 37° for

an additional 45 min. Double strand cleavage products are produced.

After quenching the reaction with EDTA (6 mM final concentration) and SDS (0.1% final concentration), the uncleaved fraction, enriched in mutation-free sequence, can be isolated by electrophoresis through non-denaturing agarose or polyacrylamide gels, depending on DNA size. The desired fragment can be isolated from the gel by methods well-known in the field (Ausubel, F.M. et al. Current Protocols in Molecular Biology, John Wiley and Sons Inc.).

Example 5: Rending molecules containing a polymerase generated mutation inert to further amplification

A population of DNA molecules is enzymatically amplified and subjected to denaturation and reannealing and MutHLS double strand cleavage as in Example 4. DNA products (about 250 ng) are incubated in a 50 μl reaction containing 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT, 50 μg/ml BSA, 25 μM of each dideoxynucleoside-5'-triphosphate (ddGTP, ddATP, ddCTP, and ddTTP), and one unit of exonuclease-free Klenow DNA polymerase at 37°C for 3 hours. The reaction is quenched by addition of EDTA to 20 mM, and extraction with phenol then ether. After removal of unincorporated ddNTPs using a silica matrix spin column (See Example 4), the resulting population of DNA molecules may then be subjected to additional rounds of PCR as required.

Example 6: Removal of molecules containing a polymerase-generated mutation utilizing MutHLS reaction followed by enzymatic degradation

A population of DNA molecules is enzymatically amplified using primers with 5'-OH termini and are subjected to denaturation and reannealing and MutHLS double strand cleavage as in Example 4. DNA products (50 ng) are then incubated in 100 μl reactions containing 67 mM glycine-NaOH (pH 9.4), 25 mM $MgCl_2$, 50 μg/ml BSA, and 2.5 units of lambda exonuclease at 37°C for 60 min. Exonuclease I (0.1 unit) is then added and reaction continued at 37°C for an additional 30 min. Reactions are terminated by addition of EDTA to 20 mM, and reactions are extracted to remove the exonucleases with phenol and then ether. The population is then subject to additional rounds of amplification as required.

Example 7: Determination of the fraction of an enzymatically amplified DNA population that contains a polymerase-generated mutation

Specific primers, PCR cycles and conditions, and the polymerase utilized are determined by the specific sequence to be amplified and the objectives of the amplification procedure based on techniques and methods familiar to those who practice the art. MutHLS reactions (single or double strand cleavage), separation of cleaved and uncleaved molecules, and quantitation can be carried

out as previously described. Other labeling, visualization, and quantitation techniques utilized with nucleic acids, that are known to those who practice the art, such as fluorescent labeling and staining with ethidium bromide, are suitable for use in this aspect of the invention.

Example 8: Determination of the Efficiency of MutHLS Treatment for Removal of Molecules Containing Mutation from PCR Product Pools

The genomes of filamentous bacteriophages are composed entirely of essential genes; however, a noncoding region between genes II and IV exists in which foreign DNA may be inserted. This region contains a cis-acting signal for packaging and orientation of DNA within bacteriophage particles, sites for the initiation and termination of DNA synthesis, and a signal for ρ-independent termination of transcription. The presence of foreign DNA segments within this region can disrupt the cis-acting elements that control replication. All filamentous phage vectors in common use therefore contain mutations in genes II or V that compensate for such disruption.

Messing and coworkers (Messing, J. New M13 Vectors for Cloning, Methods Enzymol., 101:20-78, 1983) created a series of bacteriophage M13-*lac* hybrid vectors by insertion of the regulatory sequences and the coding information for the first 146 amino acids of the *E. coli* β-galactosidase gene (*lacZ*) into the intergenic region between genes II and

IV of M13. The host cells for these M13 hybrid phage contain an F' plasmid with a β-galactosidase gene that is defective due to the fact that it encodes an enzymatically inactive polypeptide lacking amino acids 11-41. The amino terminal fragment of β-galactosidase produced in cells infected with an M13-*lac* hybrid vector associates with the defective host polypeptide to form an enzymatically active protein (this is called α-complementation). This has allowed development of a color test to distinguish vectors that have a functional as opposed to a genetically inactive β-galactosidase gene fragment. When plated on hosts carrying the appropriate F' episome, vectors encoding the wild type β-galactosidase gene fragment will form blue plaques when the medium contains the inducer IPTG and the chromogenic β-galactosidase substrate, X-GAL. By contrast, disruption of the phage *lac*Z region by mutation or by insertion of foreign DNA blocks α-complementation and results in pale blue or colorless plaques.

This principle has been exploited for the development of an assay for assessment of the replication fidelity of DNA polymerases (Benenek, K. and Kunkel, T. A. Analyzing Fidelity of DNA Polymerases. Methods Enzymol. 262:217-232, 1995). In this method an M13 substrate was constructed which contained a single-strand gap spanning the phage *lac*Z sequence. The gap was then filled by a polymerase of interest, the reaction products introduced into *E. coli* host cells that support α-complementation, and plated in the presence of X-GAL and IPTG. If the

gap-filling DNA synthesis is error-free, the β-galactosidase peptide produced will complement the defective β-galactosidase of the host, and X-GAL will be hydrolyzed in these cells to produce blue plaques. Pale blue or clear plaques are produced when errors made during the gap-filling reaction result in production of a β-galactosidase polypeptide of altered amino acid sequence that is incapable of complementation.

A variation of this fidelity assay was used to assess the efficiency of MutHLS treatment for removal of sequences containing polymerase-induced mutations that occur during the process of PCR. As outlined in Fig. 6, a region of M13mp18 spanning the β-galactosidase gene fragment was amplified using the polymerase chain reaction and products were denatured and reannealed. MutHLS treatment was then used to cleave both strands of molecules containing point mutations or small insertions and deletions, producing double-strand breaks in such molecules. Full-length products were then isolated and ligated into an M13mp18 molecule to which the region corresponding to the PCR product had been removed, thus replacing the wild type M13mp18 fragment with the corresponding PCR product. Ligation products were then transfected into *E. coli* cells and plated in the presence of IPTG, X-GAL, and an appropriate host strain. The presence of dark blue plaques indicates clones containing wild-type β-galactosidase gene fragments and pale blue or clear plaques indicates clones containing mutations in the

β-galactosidase gene. The proportion of pale blue and clear plaques to total plaques in molecules treated with MutHLS compared to the corresponding proportion in untreated molecules indicates the degree of reduction of the presence of mutant sequences.

PCR Amplification of DNAs

A 1611 base pair (bp) fragment of M13mp18 was amplified which spanned the 390 bp β-galactosidase gene segment of interest. With respect to the viral strand sequence, a *Dra* III site is located 256 bp from the 5' end of the PCR product and a *Bgl* II site is located 134 bp from the 3' end of the PCR product (see Fig. 6). PCR reactions (100 μl) contained 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 2 mM $MgCl_2$, 200 μM each dNTP (Pharmacia), 100 pmol each primer (21 nucleotides long with 5'-OH termini; Oligos Etc., Guliford, CT), 5 μg of T4 gene 32 protein (Boehringer Mannheim), 50 ng template DNA (M13mp18), and 2.5 units of AmpliTaq DNA polymerase (Perkin Elmer). The sequence of the forward primer was 5'-TTA TAC GTG CTC GTC AAA GCA-3' (SEQ. ID. NO. 10) corresponding to nucleotides 5458-5478 in M13mp18 and the sequence of the reverse primer was 5'-AAT GCC TGA GTA ATG TGT AGG-3' (SEQ. ID. NO. 11) corresponding to nucleotides 7048-7069 of M13mp18. Reactions (30 cycles) were performed using a Perkin Elmer Gene Amp 9600 thermocycler with incubations at 94°C for 15 sec, 60°C for 15 sec, and 72°C for 1 min. Products were denatured and reannealed immediately following amplification by heating

to 95°C for 1 min and incubation at 65°C for 60 min followed by incubation at 37°C for 30 min. EDTA was then added to 20 mM and reactions were extracted with phenol followed by binding to a silica matrix spin column (Pierce Xtreme DNA purification columns) and eluted with distilled H2O to remove PCR primers, dNTPs, and salts. Products were quantitated by an ethidium bromide dot method as follows: Samples (2.0 μl of an appropriate dilution) and DNAs of known concentration were added to 8 μl of 1 μg/ml ethidium bromide and spotted onto plastic wrap. UV-induced fluorescence was measured using a Photometrics cooled charge-coupled device imager. The concentration of PCR products was determined by comparison to the fluorescence of the standards.

MutHLS Reactions

Reactions (50 μl total) were assembled as follows: 20 μl of 125 mM HEPES (pH 8.0), 50 mM KCl, 2.5 mM dithiothreitol (DTT), 125 μg/ml bovine serum albumin (BSA), 5 mM ATP, 10 mM $MgCl_2$, and 1 μg PCR DNA were preincubated at 37°C for 8 min. Reactions were then initiated by adding 30 μl of a premixed solution of 5 μg MutS (Su, S.-S. and Modrich, P., Proc. Natl. Acad. Sci. U.S.A. 83, 5057-5061, 1986), 12 μg MutL (Grilley, M., Welsh, K. M., Su, S.-S. & Modrich, P., J. Biol. Chem. 264:1000-1004, 1989) and 18 ng of MutH (Welsh, K. M., Lu, A.-L., Clark, S. & Modrich, P., J. Biol. Chem. 262:15624-15629, 1987) in 20 mM potassium phosphate (pH 7.4), 50 mM KCl, 0.1 mM EDTA, 1 mM DTT, and

1 mg/ml BSA. Incubation was continued for 45 min at 37°C. An additional 30 μl of a premixed solution of 5 μg MutS, 12 μg MutL, and 18 ng MutH in the same buffer described above was then added as well as 3 μl of a 10X buffer solution containing 500 mM HEPES (pH 8.0), 200 mM KCl, 10 mM DTT, 20 mM ATP, and 40 mM $MgCl_2$. Incubation was continued at 37°C for 45 min. An additional 30 μl of a premixed solution of 5 μg MutS, 12 μg MutL, and 18 ng MutH in the same buffer as described above was then added as well as 3 μl of a 10X buffer solution containing 500 mM HEPES (pH 8.0), 200 mM KCl, 10 mM DTT, 20 mM ATP, and 40 mM $MgCl_2$. Incubation was continued at 37°C for an additional 45 min. Reactions were terminated by addition of EDTA to 10 mM, and extracted with phenol followed by extraction with ether. DNA was then bound to a Qiagen column (QIAquick spin column) and eluted with distilled H_2O to concentrate the DNA and to remove proteins, salts, and all other reaction components.

Restriction Digestion and Gel Purification

After MutHLS treatment the DNA was digested with *Dra* III and *Bgl* II in a 20 μl reaction containing 1 μg DNA, 100 μM NaCl, 50 mM Tris-HCl (pH7.5), 10 mM $MgCl_2$, 1 mM DTT, 100 μg/ml Bovine Serum Albumin, 4 units *Bgl* II (New England Biolabs), and 10 units *Dra* III (Amersham). Reaction components were incubated at 37°C for 2 hr and the reaction was terminated by addition of EDTA to 10 mM. Products were precipitated with ethanol followed by

electrophoresis through 1% agarose in 40 mM Tris-acetate, 1 mM EDTA (final pH 7.5). The full-length band (corresponding to the *Dra* III/*Bgl* II fragment of M13mp18) was recovered using a Gene Clean Kit (Bio 101) according to recommendations of the manufacturer. The recovered DNA was quantitated using the ethidium bromide dot method described above.

Ligation Reactions

The PCR product corresponding to the *Dra* III/*Bgl* II fragment of M13mp18, recovered after MutHLS treatment, *Dra* III/*Bgl* II digestion, and gel purification, was ligated into an M13mp18 derivative in which the *Dra* III/*Bgl* II fragment had been removed. The ligation reactions (20 μl) contained 50 ng PCR DNA fragment, 50 ng M13mp18 (without the *Dra* III/*Bgl* II fragment), 66 mM Tris-HCl (pH 7.6), 6.6 mM $MgCl_2$, 10 mM DTT, 66 μM ATP, and 0.5 Weiss unit of T4 DNA ligase. Reactions were incubated at 16°C for 12-16 hrs.

Transfections

Ligation reaction products were transfected into XL2-Blue Ultracompetent Cells (Stratagene) following the included protocol. An aliquot (50 μl) of the transfection reaction was added to a tube at 49°C containing 4 mls of Luria-Bertani soft agar, 4 mg of X-GAL (Amersham), and 800 μg of IPTG (Amersham). Then 200 μl of a log-phase culture of XLI-Blue was added. The soft agar mixture was poured

onto a Luria-Bertani plate and allowed to solidify. Plates were incubated at 37°C 12-16 hrs, and 3,000 to 12,000 plaques then scored according to color as mutant or wild type.

Results obtained with three independent samples of PCR-amplified DNA are summarized in Table 1. As noted above, dark blue plaques indicate the presence of a clone containing a wild type β-galactosidase gene fragment and pale blue or clear plaques indicate the presence of a clone containing a mutation somewhere within this region. This assay detects single base substitution mutations at 114 positions within the 390 bp β-galactosidase a fragment, as well as single nucleotide frameshifts at 150 positions (Eckert, K. A.. and Kunkel, T. A., Nucl. Acids Res. 18:3739-3744, 1990). As shown in Table 1, the MutHLS double strand cleavage reaction reduced the incidence of mutant plaques by 88-93%.

TABLE 1

Experiment	Mutant Plaques	Wild Type Plaques	Mutant Fraction
			%
-HLS 1	127	3,546	3.58
-HLS 2	204	5,591	3.65
-HLS 3	213	6,947	3.07
+HLS 1	23	5,191	0.44
+HLS 2	31	11,339	0.27
+HLS 3	45	11,941	0.38

Mutant Frequency = 100 x [number of mutant plaques/total number of plaques]

Other embodiments are within the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Modrich, Paul L.
Smith, Jane E.

(ii) TITLE OF INVENTION: METHODS FOR USE OF MISMATCH REPAIR SYSTEMS FOR THE DETECTION AND REMOVAL OF MUTANT SEQUENCES THAT ARISE DURING ENZYMATIC AMPLIFICATION

(iii) NUMBER OF SEQUENCES: 11

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Lyon & Lyon
(B) STREET: 633 West Fifth Street Suite 4700
(C) CITY: Los Angeles
(D) STATE: California
(E) COUNTRY: U.S.A.
(F) ZIP: 90071-2066

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: IBM P.C. DOS 5.0
(D) SOFTWARE: Word Perfect 5.1

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/008,673
(B) FILING DATE: December 15, 1995

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Warburg, Richard J.
(B) REGISTRATION NUMBER: 32,327
(C) REFERENCE/DOCKET NUMBER: 223/147

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (213) 489-1600
(B) TELEFAX: (213) 955-0440
(C) TELEX: 67-3510

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CAGAACTTTA AAAGTGCTCA T 21

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATGCAGCAAC GAGACGTCAC G 21

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GATAAGAGGT CATTTTTGCG G 21

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AGACCGGAAG CAAACTCCAA C 21

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GCCCGAAAGA CTTCAAATAT C 21

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TTATAGTCAG AAGCAAAGCG G 21

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGATAGCGTC CAATACTGCG G 21

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

ATCATAACCC TCGTTTACCA G 21

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CCAGCAAGGC CGATAGTTTG A 21

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TTATACGTGC TCGTCAAAGC A 21

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

AATGCCTGAG TAATGTGTAG G 21

WHAT IS CLAIMED IS:

1. Method for removing DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprising the steps of:

denaturing and reannealing said population of DNA duplexes,

contacting said reannealed DNA duplexes with a mismatch repair system such that each strand is cleaved in DNA duplexes containing a base pair mismatch, and

separating said cleaved DNA duplexes from uncleaved DNA duplexes.

2. The method of claim 1, wherein said separating is by electrophoresis through a gel.

3. Method for removing DNA molecules containing one or more polymerase-generated mutations in a population of DNA duplexes that have been enzymatically amplified using primers containing 5'-hydroxyl termini comprising the steps of:

denaturing and reannealing said population of DNA duplexes,

contacting said reannealed DNA duplexes with a mismatch repair system such that each strand in DNA duplexes containing a base pair mismatch is cleaved, and

further contacting said population of reannealed DNA duplexes with exonucleases so that said DNA duplexes containing a base pair mismatch are enzymatically degraded.

4. The method of claim 1 or 3, wherein said mismatch repair system comprises components of the methyl-directed mismatch repair system of *E. coli* and includes the MutS, MutL, and MutH proteins.

5. Method for rending inert to further amplification DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprising the steps of:

denaturing and reannealing said population of DNA duplexes,

contacting said reannealed DNA duplexes with a mismatch repair system such that each strand is cleaved in DNA duplexes containing a base pair mismatch, and,

further contacting said cleaved DNA duplexes with dideoxynucleoside-5'-triphosphates and a DNA polymerase.

6. The method of claim 5, wherein said mismatch repair system comprises components of the methyl-directed mismatch repair system of *E. coli* and includes the MutS, MutL, and MutH proteins and the dideoxynucleoside-5'-triphosphate is dideoxyguanosine-5'-triphosphate.

7. Method for determining the fraction of an enzymatically amplified DNA population that contains polymerase-generated mutations comprising the steps of:

denaturing and reannealing said population of DNA duplexes,

contacting said reannealed DNA duplexes with a mismatch repair system such that at least one strand is cleaved in DNA duplexes containing a base pair mismatch,

separating cleaved DNA duplexes from uncleaved DNA duplexes, and

determining the fraction of cleaved DNA duplexes relative to uncleaved DNA duplexes as an indication of the fraction of enzymatically amplified DNA that contain polymerase-generated mutations.

8. The method of claim 7, wherein said mismatch repair system comprises components of the methyl-directed mismatch repair system of *E. coli* and includes the MutS, MutL, and MutH proteins.

9. Method for detecting the presence of DNA polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprising the steps of:

denaturing and reannealing said population of DNA duplexes,

contacting said reannealed DNA duplexes with a mismatch repair system under conditions such that a duplex

containing a polymerase generated mutation is modified by the introduction of an endonucleolytic incision in at least one strand of said duplex, and

detecting the product of said endonucleolytic incision as an indication of the presence of polymerase generated mutations.

10. The method of claim 9 wherein the detection of the product of said endonucleolytic incision is by altered electrophoretic mobility under denaturing conditions.

11. The method of claim 9, wherein said mismatch repair system comprises components of the methyl-directed mismatch repair system of *E. coli* and includes the MutS, MutL, and MutH proteins.

12. Method for detecting the presence of DNA polymerase-generated mutations in a population of enzymatically amplified DNA duplexes produced from DNA duplexes lacking a sequence subject to mismatch provoked endonucleolytic cleavage comprising the steps of:

enzymatically amplifying a population of DNA molecules utilizing primers containing a sequence subject to mismatch provoked endonucleolytic cleavage,

denaturing and reannealing said population of DNA duplexes,

contacting said reannealed DNA duplexes with a mismatch repair system under conditions such that an endonucleolytic incision is introduced in at least one strand of a duplex containing a polymerase generated mutation, and

detecting the product of said endonucleolytic incision as an indication of the presence of polymerase generated mutations.

13. The method of claim 12 wherein said sequence subject to mismatch provoked endonucleolytic cleavage is a d(GATC) site.

14. The method of claim 12 wherein the detection of the product of said endonucleolytic incision is by altered electrophoretic mobility under denaturing conditions.

15. The method of claim 12, wherein said mismatch repair system comprises components of the methyl-directed mismatch repair system of *E. coli* and includes the MutS, MutL, and MutH proteins.

16. Kit for amplifying DNA molecules and for removing DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprising:

primers with 5'-hydroxyl termini,

components of a mismatch repair system, and

exonucleases.

17. Kit for rendering inert to further amplification DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprising:

components of a mismatch repair system,

a dideoxynucleoside-5'-triphosphate, and

a DNA polymerase.

18. Kit for amplifying DNA molecules lacking a sequence subject to mismatch provoked endonucleolytic cleavage and detecting or removing DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprising:

primers containing sequence subject to mismatch provoked endonucleolytic cleavage, and

components of a mismatch repair system.

19. Kit for removing DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes comprising:

components of a mismatch repair system, and

exonucleases.

20. Method for removing DNA molecules containing one or more polymerase-generated mutations in a population of enzymatically amplified DNA duplexes produced from DNA duplexes lacking a sequence subject to mismatch provoked endonucleolytic cleavage comprising the steps of:

enzymatically amplifying a population of DNA molecules utilizing primers containing a sequence subject to mismatch provoked endonucleolytic cleavage,

denaturing and reannealing said population of DNA duplexes,

contacting said reannealed DNA duplexes with a mismatch repair system under conditions such that each strand is cleaved in a DNA duplex containing a polymerase generated mutation, and

separating said cleaved DNA duplexes from uncleaved DNA duplexes.

FIG. 1.

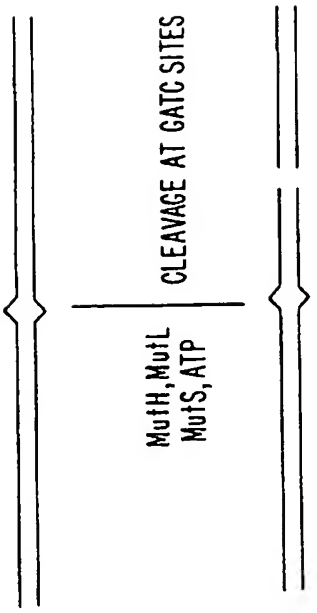

MutH, MutL
MutS, ATP
CLEAVAGE AT GATC SITES

FIG. 2.

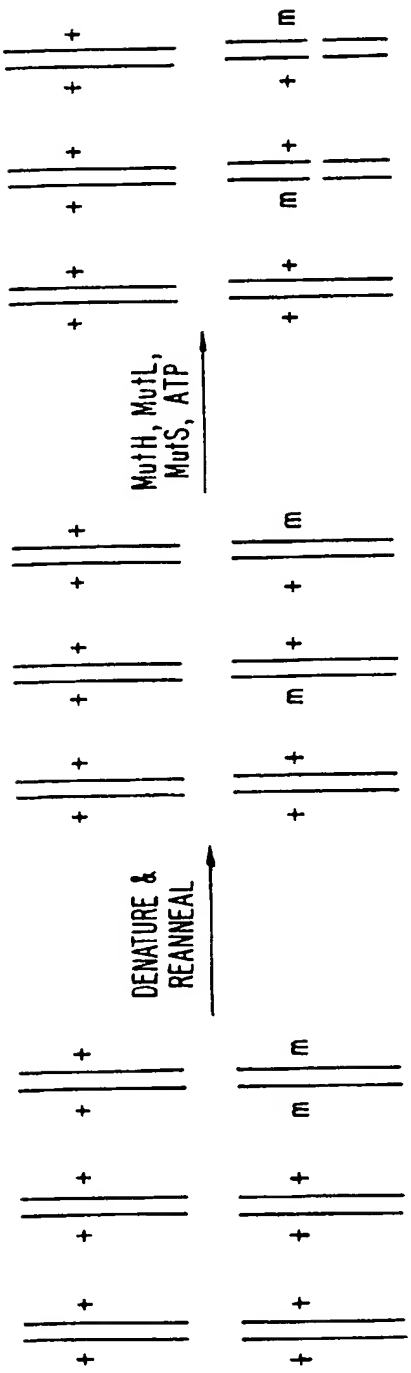

MutH, MutL,
MutS, ATP
DENATURE &
REANNEAL

FIG. 3A.

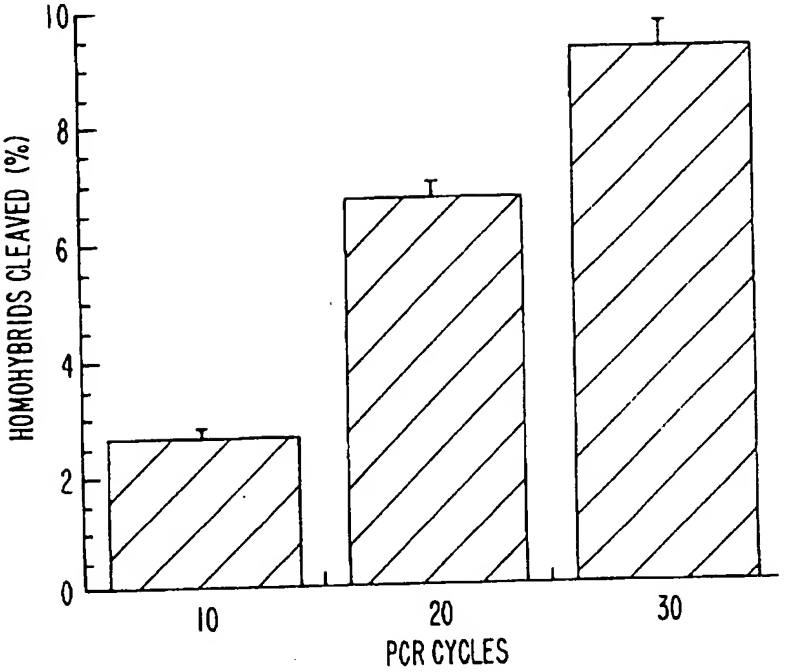
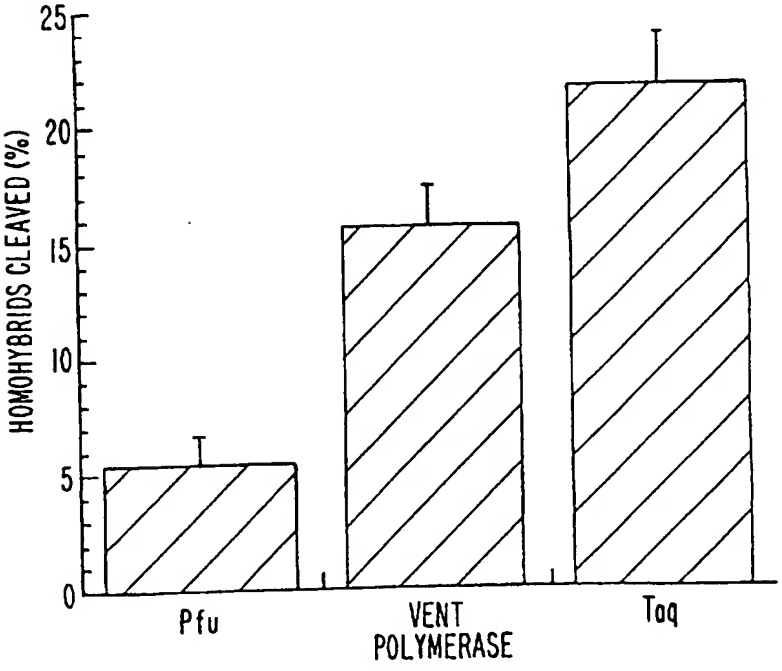

HOMOHYBRIDS CLEAVED (%)
0
2
4
6
8
10
10
20
30
PCR CYCLES

FIG. 3B.


HOMOHYBRIDS CLEAVED (%)
0
5
10
15
20
25
Pfu
VENT
POLYMERASE
Taq

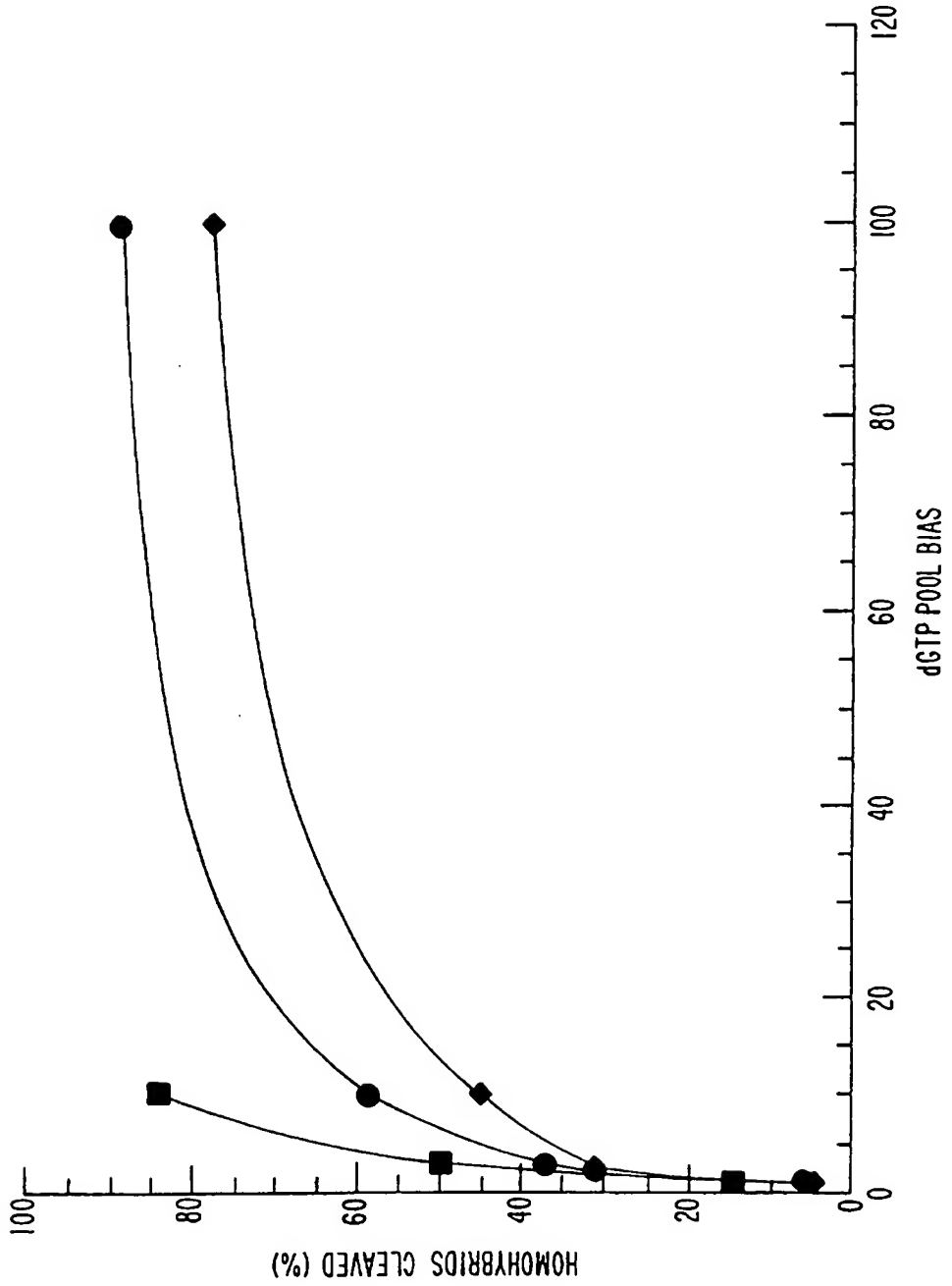

HOMOHYBRIDS CLEAVED (%)
100
80
60
40
20
0
dGTP POOL BIAS
0
20
40
60
80
100
120

FIG. 4.

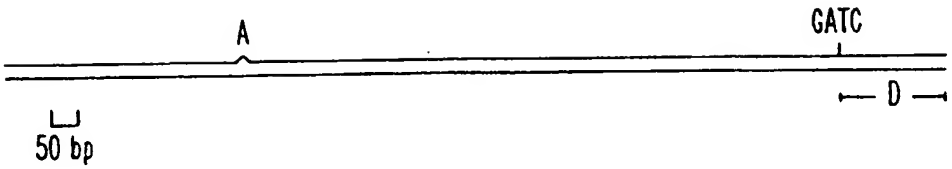

A
GATC
D
50 bp

FIG. 5A.

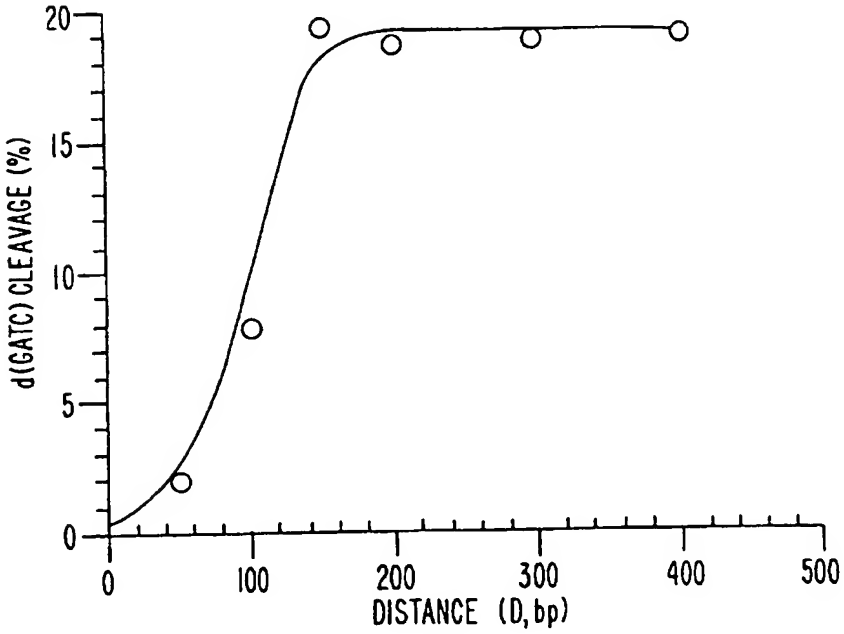
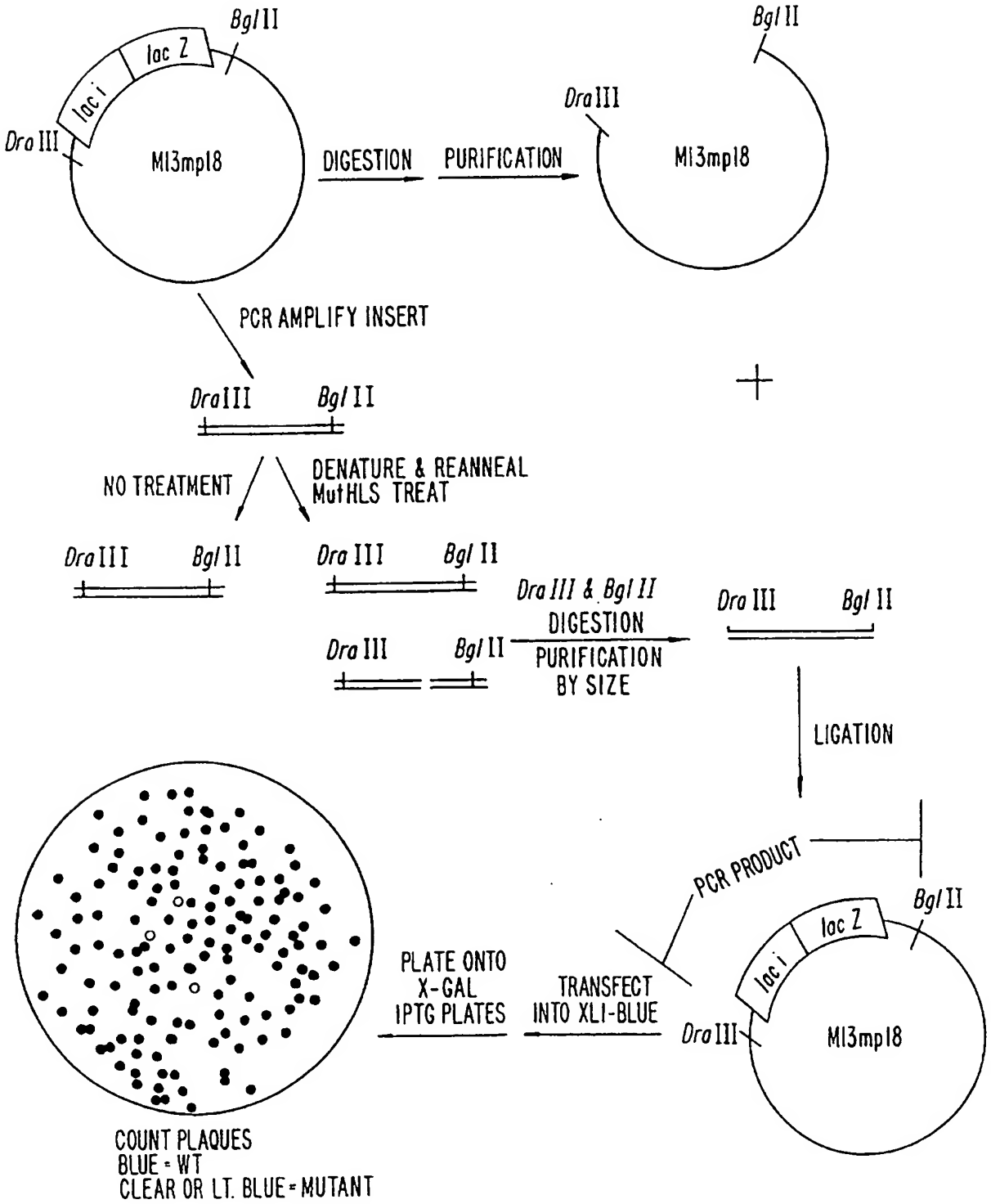

d(GATC) CLEAVAGE (%)
20
15
10
5
0
0
100
200
300
400
500
DISTANCE (D, bp)

FIG. 5B.

FIG. 6.


Bgl II
lac Z
lac i
Dra III
MI3mpI8
DIGESTION
PURIFICATION
Dra III
MI3mpI8
Bgl II
PCR AMPLIFY INSERT
Dra III
Bgl II
NO TREATMENT
DENATURE & REANNEAL
MutHLS TREAT
Dra III
Bgl II
Dra III
Bgl II
Dra III & Bgl II
DIGESTION
PURIFICATION
BY SIZE
Dra III
Bgl II
Dra III
Bgl II
LIGATION
PCR PRODUCT
Bgl II
lac Z
lac i
Dra III
MI3mpI8
PLATE ONTO
X-GAL
IPTG PLATES
TRANSFECT
INTO XLI-BLUE
COUNT PLAQUES
BLUE = WT
CLEAR OR LT. BLUE = MUTANT

INTERNATIONAL SEARCH REPORT

Internat Application No
PCT/US 96/20075

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 95 12689 A (GENE CHECK INC ;WAGNER ROBERT E JR (US)) 11 May 1995 see page 21, line 9 - line 25; claim 11 ---	1-20
Y	WO 95 12688 A (US BIOCHEMICAL CORP) 11 May 1995 cited in the application see whole document, especially page 49, ln 30-page 50, ln 18 page 16; lns 15-33; claim 56 ---	1-20
Y	WO 95 29258 A (ST JAMES AND SEACROFT UNIVERSI ;TAYLOR GRAHAM ROY (GB)) 2 November 1995 see the whole document --- -/--	1-4,7,8, 16,19

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

3

Date of the actual completion of the international search	Date of mailing of the international search report
24 April 1997	14.05.97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Osborne, H

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Internat Application No
PCT/US 96/20075

C.(Conunuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 95 07361 A (PASTEUR INSTITUT ;INST NAT SANTE RECH MED (FR); MEO TOMMASO (FR);) 16 March 1995 see claims 1-8	1-4,7,8, 16,19
Y	NATURE GENETICS, vol. 9, 1 February 1995, pages 177-183, XP000600255 MASHAL R D ET AL: "DETECTION OF MUTATIONS BY CLEAVAGE OF DNA HETERODUPLEXES WITH BACTERIOPHAGE RESOLVASES" see the whole document	1-20
Y	NUCLEIC ACIDS RESEARCH, vol. 23, no. 19, October 1995, pages 3944-48, XP002030020 WAGNER R ET AL: "Mutation detection using immobilized binding protein (MutS)" see page 3944, paragraph 2 see page 3947, paragraph 2 - page 3948, paragraph 2	1-20
P,X	PROC. NATL. ACAD. SCI. USA, vol. 93, April 1996, pages 4374-4379, XP002030021 SMITH ET AL: "Mutation detection with MutH, MutL, and MutS mismatch repair proteins" see the whole document	1-20
A	BIOTECHNIQUES, vol. 15, 1993, pages 489-501, XP000389938 TESSIER D ET AL: "PCR-assisted large insertion/deletion mutagenesis" see page 498, paragraph 1 - paragraph 3	12,18,20
A	EP 0 593 789 A (SUMITOMO METAL IND) 27 April 1994 see claims 1,2	12,18,20

3

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 96/20075

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9512689 A	11-05-95	AU 1050895 A	23-05-95
		CA 2175222 A	11-05-95
		EP 0733126 A	25-09-96
WO 9512688 A	11-05-95	US 5556750 A	17-09-96
		US 5459039 A	17-10-95
		CA 2175490 A	11-05-95
WO 9529258 A	02-11-95	AU 2315195 A	16-11-95
		EP 0758402 A	19-02-97
WO 9507361 A	16-03-95	FR 2709761 A	17-03-95
		CA 2171469 A	16-03-95
		EP 0717781 A	26-06-96
EP 0593789 A	27-04-94	JP 5308999 A	22-11-93
		WO 9323567 A	25-11-93